Effect of correcting outcome data for case mix: an example from stroke medicine

Richard J Davenport, Martin S Dennis, Charles P Warlow

Abstract

Objective—To show the influence of variations in case mix on clinical outcome indicators for patients admitted to hospital with acute stroke.

Design—"Before and after" cohort study, with prospective, consecutive identification of patients and prospective follow up; multiple logistic regression analyses to correct for case mix variations.

Setting—University teaching hospital.

Subjects—216 patients with stroke identified before the introduction of an organised stroke service, and 252 patients with stroke identified after its introduction.

Main outcome measures—Case fatality at 30 days and 12 months; for survivors at 12 months, proportions of patients who were independent (according to the Oxford handicap scale) and of those living at home.

Results—Crude outcome data suggested that patients in the cohort identified after the introduction of the stroke service were significantly more likely to be alive, independent, and living at home than patients managed before the stroke service. After adjustment for age and sex these "improvements" were less impressive but still significant. After adjustment for many other possible prognostic indicators, however, the differences between the two groups for all four outcomes were non-significant, suggesting that the "improvements" may have been entirely due to differences in case mix between the two cohorts, rather than the new stroke service.

Conclusions—Variations in case mix have a crucial influence on the interpretation of outcome data, and this is particularly important in non-randomised comparative studies. Such studies, comparing performance within and between different provider units, are likely to become increasingly common in the new reformed NHS. To allow meaningful interpretation, these studies must try to correct for case mix.

Introduction

The confounding influence that variations in case mix may exert on clinical outcome is well recognised.¹⁻³ Despite this, the government has published for Scottish hospitals clinical outcome indicators for several diseases—including stroke—that have only been corrected for age and sex.⁴ These data have been widely interpreted by the media (and even by some doctors and managers) as league tables of hospital performance. To assess the impact of case mix on the outcome after acute stroke, we compared the outcome of patients managed before and after the introduction of an organised stroke service in our hospital. This service included an inpatient stroke unit, an intervention that has been shown in a systematic review of randomised controlled trials to reduce case fatality and dependence.⁵

Methods

We prospectively identified consecutive stroke patients who needed inpatient management in our hospital (excluding those with subarachnoid haemorrhage and those admitted to the regional neurosciences unit) by daily consultation with the admitting medical and nursing teams and perusal of admissions records. All patients were assessed by a stroke physician who completed a standard data form; the data were then transferred to an electronic database (stroke register). We followed up survivors either with a face to face or telephone interview or with a postal questionnaire. We measured four important outcomes: case fatality at 30 days and 12 months; the proportion of survivors living at home at 12 months after their stroke; and the proportion of survivors who were independent (<3 on the Oxford handicap scale⁶) at 12 months. We calculated uncorrected odds ratios and 95% confidence intervals for these outcomes. We then performed multiple logistic regression, forcing in 19 different indicators of case mix (box); we chose these variables on the basis of likely clinical importance and from our previous experience with prognostic models derived from large community and hospital based stroke registers (C Counsell, personal communication). This provided corrected odds ratios for the outcomes. We also corrected the data for age and sex alone.

Results

We identified 216 strokes before and 252 strokes after the introduction of the stroke unit, over a period of 27 months. Three patients were lost to follow up at 30 days and four at 12 months; all were overseas visitors. Independence data for six survivors were unavailable (two in the cohort managed before the introduction of the stroke unit and four in the cohort managed after). Figure 1 shows the uncorrected odds ratios; patients admitted after the introduction of the stroke unit had very significantly better outcomes—for example, a 50% increase in the odds of living at home at 12 months. Correcting for age and sex alone reduced these estimates of treatment effect, but they remained significant. When the data were corrected for case mix, however, all the estimates became non-significant, with odds ratios approaching unity.

Discussion

Because our study was a non-randomised, "before and after" comparison, we were aware that variations in case mix might significantly bias the outcomes. We therefore corrected the crude outcomes using multiple logistic regression analysis, which is a method of simultaneously adjusting for the effects of several different variables. Before discussing the implications of our results, we consider some of the problems of using these complicated mathematical models.⁸

Firstly, such models may "overcorrect" the data—that is, they may obscure a real but moderate treatment

University of Edinburgh,
Department of Clinical
Neurosciences, Western
General Hospital,
Edinburgh EH4 2XU
Richard J Davenport, *clinical
research fellow*
Martin S Dennis, *senior
lecturer in stroke medicine*
Charles P Warlow, *professor of
medical neurology*

Correspondence to:
Dr Davenport.

BMJ 1996;312:1503-5

Variables forced into multiple logistic regression modelling (measured at time of stroke)

Variable	Definition
Age (years)	Not applicable
Able to lift both arms above gravity	Not applicable
Diabetes mellitus	Previous history recorded
Employed	Employed at time of stroke
Assessed for stroke register within 48 hours of onset	Not applicable
Normal eye opening	Glasgow coma eye score 4
Normal motor score	Glasgow coma motor score 6
Normal verbal score	Glasgow coma verbal score 5
Haemorrhagic stroke	Haemorrhage on computed tomography, magnetic resonance imaging, or postmortem
Prestroke independence	Oxford handicap scale <3
High systolic blood pressure	>159 mm Hg recorded since onset of stroke
Low systolic blood pressure	<121 mm Hg recorded since onset of stroke
Living alone	Not applicable
Cohort	Before or after stroke service was available
Previous myocardial infarction	History of previous myocardial infarction from patient, general practitioner, or hospital notes
Sex	Not applicable
Total anterior circulation infarct (assumed or definite)	According to Oxfordshire Community Stroke Project classification ⁷
Unable to walk	Not applicable
Urinary incontinence within 7 days of onset	More than two episodes of incontinence or catheterised since onset of stroke service (recorded in nursing or hospital notes)

effect by including too many variables. Wasson *et al* suggested that for every variable there should be at least five patients with the outcome of particular interest.⁹ In our models we used 19 variables, and there were considerably more than 95 outcome events for each outcome we measured.

Secondly, models derived from one dataset may not perform well on another, independent dataset. In our department we have developed prognostic models for a variety of outcomes, including the four in the present study, which rely on five or six variables (C Counsell and M McDowall, personal communication). These models were derived from a community based stroke register (the Oxfordshire Community Stroke Project¹⁰) and have been validated on two further independent datasets (one community and one hospital based register). When we applied these validated models to the results of the current study we obtained almost identical corrected odds ratios as those in the figure, which was reassuring.

IMPLICATIONS OF FINDINGS

Our results confirm the dangers of comparing outcome measures that have not been adjusted for case mix; it was case mix, rather than a beneficial effect of our stroke unit,

that seemed to be responsible for much of the observed "improvements" in outcome (although because of the imprecision of the corrected data, the results are still consistent with a moderate but non-significant beneficial effect). Several factors may have explained why case mix changed, but one factor was the closure of our hospital's casualty department during the study. Similar findings have been shown for other diseases.¹¹⁻¹³ While epidemiologists and clinical trialists have been aware of this problem for some time, its importance in everyday clinical practice is perhaps less readily acknowledged.

Although randomised controlled trials remain the gold standard for assessing interventions, we believe that there are two areas where non-randomised comparative studies are likely to become increasingly common. Firstly, individual provider units are increasingly aware of the need to provide evidence of the quality of care that they deliver; because randomised controlled trials within single units are impractical and sometimes unethical, often the most feasible design for these studies is a "before and after" one. These non-randomised studies are susceptible to several biases, with the confounding effect of case mix variation one of the most potent; therefore, if providers wish to draw valid comparisons between outcomes, they must collect potentially important case mix variables as well as outcome measures. Secondly, the new NHS reforms encourage competition between provider units, with purchasers empowered to select the best treatment packages for their patients. Currently, little information is available to purchasers on which to base such decisions, as there are few comparative data between different hospitals for similar conditions.

Recently, the government published limited outcome indicators for several diseases, comparing the performance of hospitals in Scotland.⁴ In the foreword to this document the problems of case mix were acknowledged, and it was stated that the purpose of publication was not to provide an indication of the best and worst hospitals for general practitioners and patients but to improve overall standards of care. It was added that trusts that underperformed should immediately review their treatments in the relevant areas, and health boards and purchasers should be encouraged to question the quality of care delivered in these centres; yet these suggestions may be entirely unjustified as the data took no account of case mix. We have clearly shown that if the government and other agencies wish to provide meaningful comparative outcome data they must address the question of case mix before publishing any more potentially unreliable figures.

Addendum

After we submitted this paper for publication the government published an updated version of their outcome

Fig 1—Odds ratios (95% confidence intervals) comparing standard ward care with stroke service care; data for four outcomes before and after correction for age and sex only, and for 19 indicators of case mix

Key messages

- In the reorganised NHS non-randomised studies comparing performance both between and within provider units will become increasingly common
- Variations in case mix wield a vital influence over outcome
- The government has published clinical outcome indicators for several diseases in Scottish hospitals that have not been adequately corrected for case mix
- Interpretation of outcome data from non-randomised studies may be confounded by case mix, and researchers must try to adjust their data for case mix variables to allow meaningful comparisons

indicators.¹⁴ In addition to correction for age and sex, these data were also corrected for social deprivation with the Carstairs index,¹⁵ based on postcode sectors of residence according to the 1991 census. Data were also corrected for pre-existing morbidity, which was based on the principal Scottish Morbidity Record 1 (SMR1) diagnoses for the preceding five years, not on secondary diagnoses recorded at the time of discharge after stroke—that is, there was no correction for other conditions that had not been noted in previous hospital admissions. It was acknowledged that this system represented an extremely crude and only partial adjustment. Moreover, there was still no correction for crucially important case mix variables for stroke prognosis, such as level of consciousness on admission.

We are grateful to Mr Jim Slattery for advice regarding the statistical methods used in this paper.

Funding: RJD is funded by the Medical Research Council (UK); MSD and the stroke register were funded by the Stroke Association (UK).

Conflict of interest: None.

- 1 Orchard C. Comparing health care outcomes. *BMJ* 1994;308:1493-6.
- 2 Green J, Wintfield N, Sharkey P, Passman LJ. The importance of severity of illness in assessing hospital mortality. *JAMA* 1990;263:241-6.
- 3 Bion J. Outcomes in intensive care. *BMJ* 1993;307:953-4.
- 4 Scottish Office. *Clinical outcome indicators*. Edinburgh: HMSO, 1994.

- 5 Stroke Unit Trialists' Collaboration. A systematic review of specialist multi-disciplinary team (stroke unit) care for stroke inpatients. In: Warlow C, Van Gijn J, Sandercock P, eds. *Stroke module of the Cochrane database of systematic reviews*. London: BMJ Publishing, 1995.
- 6 Bamford JM, Sandercock PAG, Warlow CP, Slattery J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1989;20:828.
- 7 Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet* 1991;337:1521-6.
- 8 Wyatt JC, Altman DG. Prognostic models: clinically useful or quickly forgotten? *BMJ* 1995;311:1539-41.
- 9 Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. *N Engl J Med* 1985;313:793-9.
- 10 Bamford J, Sandercock P, Dennis M, Warlow C, Jones L, McPherson K, et al. A prospective study of acute cerebrovascular disease in the community: the Oxfordshire Community Stroke Project 1981-86. 1. Methodology, demography and incident cases of first-ever stroke. *J Neurol Neurosurg Psychiatry* 1988;51:1373-80.
- 11 Rockall TA, Logan RFA, Devlin HB, Northfield TC, for the National Audit of Acute Upper Gastrointestinal Haemorrhage. Variation in outcome after acute upper gastrointestinal haemorrhage. *Lancet* 1995;346:346-50.
- 12 Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's APACHE II study in Britain and Ireland. I. Variations in case mix of adult admissions to general intensive care units and impact on outcome. *BMJ* 1993;307:972-7.
- 13 Keeler EB, Kahn KL, Draper D, Sherwood MJ, Rubenstein LV, Reinisch EJ, et al. Changes in sickness at admission following the introduction of the prospective payment system. *JAMA* 1990;264:1962-8.
- 14 Scottish Office. *Clinical outcome indicators*. Edinburgh: HMSO, 1995.
- 15 Carstairs V, Morris R. *Deprivation and health*. Aberdeen: Aberdeen University Press, 1991.

(Accepted 3 April 1996)

Psychological complications after stillbirth—influence of memories and immediate management: population based study

Ingela Rådestad, Gunnar Steineck, Conny Nordin, Berit Sjögren

Abstract

Objective—To identify factors that may predict long term psychological complications among women who have had a stillborn child.

Design—Nationwide population based study using epidemiological methods.

Subjects—380 subjects and 379 controls who had had a stillborn or non-deformed live child in Sweden in 1991.

Results—Information was provided by 636 (84%) women. The ratio (95% confidence interval) of proportions of women with symptoms related to anxiety above the 90th centile for women who had had a stillborn child compared with those who had not was 2.1 (1.2 to 3.9). An interval of 25 hours or more from the diagnosis of death in utero to the start of delivery gave a ratio of 4.8 (1.5 to 15.9). The ratio was 2.3 (1.1 to 5.3) for not seeing the child as long as the mother had wished and 3.1 (1.6 to 6.0) for no possession of a token of remembrance.

Conclusion—It is advisable to induce the delivery as soon as feasible after the diagnosis of death in utero. A calm environment for the woman to spend as much time as she wants with her stillborn child is beneficial, and tokens of remembrance should be collected.

Introduction

It is psychologically traumatic to give birth to a stillborn child. A stillbirth is often unexpected and happens quickly, and the emotional changes experienced by the parents are enormous. Investigations have reported that 20-30% of women with perinatal loss of a child have appreciable psychiatric long term morbidity.¹⁻⁶ The rapidly changing practice in maternal care during recent years may have altered the prerequisites for these mothers being able to cope with the trauma.⁷⁻¹⁶ Twenty years ago a stillbirth was normally regarded as a "non-event."¹⁷ Today, the approach is the reverse; it is

believed that confronting parents with the reality facilitates healthy mourning.¹⁸

The new routines may have drawbacks. Bourne and Lewis argue that there now is a danger that staff may inflexibly apply the dogma, demanding that every woman should obediently inspect and hold her stillborn child.¹⁹ Leon states that checklists and behavioural protocols may result in disturbing "institutionalisation of bereavement."²⁰ To be able to disentangle all the suggested actions and divergent kinds of advice, data are needed.

Methods

We identified all mothers who had had a stillborn child in Sweden in 1991 (subjects) through the medical birth register of the National Board of Health and Welfare. In 1991, 124 201 children were born in Sweden, of whom 464 (3.7 per 1000) were stillborn (defined as a fetus at ≥ 28 weeks' gestation). Each control had delivered a live child with no deformities at the same hospital as a subject on a date as close as possible to the birth date of the corresponding subject's stillborn child. Overall, 380 subjects and 379 controls fulfilled the criteria of having given birth to a single child (stillborn or alive) and being fluent enough in Swedish to answer a questionnaire. We collected data by a postal questionnaire and safeguarded the anonymity of the investigation by letting the women return the questionnaire separately from a second form, in which they confirmed that they had replied. We collected the information during October 1994.

We assessed anxiety related and depression related symptoms using the trait anxiety inventory²¹ and the Center for Epidemiological Studies depression scale.²² We summarised the responses to the questionnaire by giving a score on a scale of 1-4 (1=least severe symptoms, 4=most severe symptoms) for each answer and dividing the sum of these scores by the number of questions answered.

In the analysis we formed dichotomised variables. For the summary average score on the trait anxiety

Centre for Caring Sciences North, Karolinska Institute, Stockholm, Sweden
Ingela Rådestad, midwife

Unit of Cancer Epidemiology, Radiumhemmet, Karolinska Institute
Gunnar Steineck, associate professor

Department of Women's and Children's Health, Karolinska Institute
Berit Sjögren, associate professor

Department of Psychiatry, University Hospital, Linköping, Sweden
Conny Nordin, associate professor

Correspondence to:
Dr I Rådestad, Centre for Caring Sciences North, Borgmästarvillan, Karolinska Sjukhuset, S-171 76 Stockholm, Sweden.

BMJ 1996;312:1505-8